SPATIAL ADAPTIVE WAVELET THRESHOLDING FOR IMAGE DENOISING

S. Grace Chang

Dept. of Electrical Engineering & Computer Science
University of California, Berkeley
CA 94720, USA
grchang@eecs.berkeley.edu

Martin Vetterli

Département d'Electricité
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
vetterli@de.epfl.ch

ABSTRACT

Wavelet thresholding with uniform threshold has shown some success in denoising. For images, we propose that this can be improved by adjusting thresholds spatially, based on the rationale that detailed regions such as edges and textures tolerate some noise but not blurring, whereas smooth regions tolerate blurring but not noise. The proposed algorithm is based on multiscale edge detection and image segmentation and then thresholding the coefficients of different regions with adaptive thresholds.

1. INTRODUCTION

When filtering random noise from an image, the two main concerns are how much of the noise granularity has been removed, and how well the edges are preserved (without blurring). Typical denoising methods are based on smoothing and stems from the notion that for a large class of images, the signal energy are compacted into a few transform coefficients and that noise contribute to the high frequency and insignificant coefficients. For Fourier based methods, the standard one is the Wiener filter, which attenuates the high frequency part of the spectrum, but as a result, removes some of the image details as well. Alternatively, the technique of suppressing coefficients in the wavelet transform domain has shown promise, where the localized nature of the coefficients makes it more suitable for locally adaptive image processing.

One method that has received considerable attention in recent years is *wavelet thresholding*, an idea that noise is removed by killing coefficients that are insignificant relative to some threshold, and whose attractiveness is due to its simplicity and effectiveness. On the theoretical side, Donoho and Johnstone [1] have shown this technique to possess some asymptotic minimax optimality properties. For image applications, it has shown success for various types of noise such as white random noise and compression artifacts [2, 6]. An extreme case of eliminating coefficients is the detection of multiscale edges [3], where all but the coefficients corresponding to edges are eliminated, a sensible idea when the noise power is so large as to render all the image details irrecoverable.

The aforementioned thresholding techniques in [1, 2, 6] rely on the likelihood that significant coefficients are due to the signal, while insignificant coefficients are due to noise, and there is no spatial discrimination to such processing (i.e. same threshold throughout). However, one can improve on this by observing that noise has varied visibility in different regions, and thus it makes sense to adapt to the changing spatial characteristics. That is, in a smooth region, noise is more visible, whereas in a detailed region such as edges and textures, one rather tolerates a little noise to keep the signal details. Hence, the threshold should be adapted in a spatial manner.

This spatial adaptive idea is illustrated in Figure 1 for a 1-D signal, where a step edge is corrupted by *iid* Gaussian additive noise, and it is denoised by ***soft-thresholding*** ('shrink' or 'kill') the wavelet coefficients with a different threshold at each scale. Figure 1 (a) shows the wavelet decomposition of the signal (3 scales and a residual). The peaks induced by the step shows strong correlation across scales, whereas the peaks due to noise do not. If a small threshold is used, the sharp edge is retained but the signal is still noisy. On the other hand, if the threshold is very large, then the reconstructed waveform is smoothed out at the step. Hence, to both retain the step and remove the noise, a small threshold is used on the step-induced peaks, and a large threshold is used elsewhere, and the result is a well denoised step function (see Figure 1 (b)). This example illustrates that a spatially adaptive threshold is a good idea. One may ask why ***hard-thresholding***, a 'keep' or 'kill' strategy, is not used. The reason is that for images, soft-thresholding usually yields more visually pleasing images than hard-thresholding, which tends to produce more spurious blips. Hence, soft-thresholding is employed in our denoising algorithm.

The basis of our algorithm is to first classify the image into different regions, followed by soft-thresholding with threshold adapted to each region. The classification consists of two parts. First, edges are detected using multiscale edge analysis described in [3, 4]. Then the remaining points are classified further into texture or smooth regions using a standard image segmentation technique.

2. DENOISING ALGORITHM

The denoising algorithm is based on soft-thresholding the wavelet coefficients, with the thresholds adapting to the different spatial characteristics (edges, textures, or smooth area). For the region classification, first the edge regions are detected from the non-edges. Then the non-edge regions are further classified into a textured or a non-textured area.

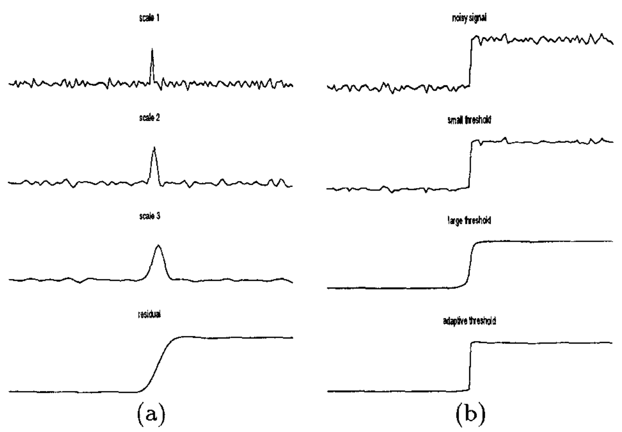

Figure 1: Motivation for using adaptive thresholding. (a) Wavelet decomposition (3 scales and the low frequency residual) of a step function corrupted by noise. (b) The original noisy step function, and the reconstructions from small and large uniform thresholding, and adaptive thresholding.

2.1. Region Segmentation

2.1.1. Edges

The method of multiscale edge detection described in [3, 4] is used to find the edges. The main ideas will be summarized here, and the reader is referred to [3, 4] for more details. With a certain class of wavelets, the nonsubsampled wavelet decomposition essentially implement the discretized gradient of the image smoothed at different scales. From traditional edge detection, one knows that the points of sharp variations occur at local maxima (called *modulus maxima*) of the gradient norm in the direction of the gradient . An isolated singularity (i.e. point of sharp variation) induces modulus maxima which propagate across scales, and this evolution can be characterized by the local Lipschitz regularity. That is, if the function f is Lipschitz α at point (x_0, y_0), then for (x, y) in its neighborhood,

$$M_s f(x, y) \leq K s^{\alpha},$$

where $M_s f$ denotes the modulus of the wavelet transform at scale s, and K is a constant. On the other hand, a noise sequence is almost everywhere singular, and its induced modulus maxima do not show coherence across scales (see Figure 1 (a)). This property allows one to detect edge points by associating local maxima points across scales.

To associate the chain of modulus maxima across scales which correspond to a singular point, the following ad hoc method is used for computational reasons. Since the first scale is very noisy, one starts the association from the second scale coefficients. For each modulus maximum larger than some threshold, if there is a modulus maximum along the same direction within a small neighborhood of the next scale, and if their ratio corresponds to $\alpha \in [0, 2]$, then they belong to the same chain. The modulus maxima of the first scale is assigned to be at the same locations as the modulus maxima of the second scale.

The chain of modulus maxima are labeled as EDGE pixels. Furthermore, since an edge induces a "hump" in the wavelet domain covering a small neighborhood (and this hump grows larger as the scale becomes coarser), the modulus maximum along with a small neighborhood around it are labeled EDGE points as well.

2.1.2. Textures and Smooth Areas

After determining the edge regions, one needs to classify the remaining pixels to texture or smooth areas. Such a classification resembles image segmentation, and has more success in the space domain rather than the transform domain. The segmentation is done by calculating a variance image, and then thresholding it between certain amplitudes [5]. Each pixel of the variance image is defined as

$$V[i, j] = \sum_{m,n \in \mathbf{C}} (f[i, j] - f[m, n])^2$$

where $f[i, j]$ is the intensity of the discretized image, and $\mathbf{C}$ is the neighborhood around pixel location $[i, j]$ (a 5×5 window is used here). To determine the thresholds, note that the variance image is roughly an estimate of the local variance, and in a smooth region contaminated by noise, it should be approximately equal to the noise variance σ^2. Hence the pixels $V[i, j]$ which satisfy the following condition are labeled SMOOTH: $1/(1 + \delta) \leq V[i, j]/\sigma^2 \leq (1 + \delta)$, where $\delta = 0.2$ is used. The remaining points are classified as TEXTURE pixels.

Such a simple segmentation of course is quite prone to error, and for example, among a patch of predominantly SMOOTH pixels, there are many holes with TEXTURE pixels, and vice versa. These holes can be closed by combinations of standard binary image operations called *erosion* and *dilation* which belong to a class of operations collectively described as morphological operations [5]. The erosion operation turns OFF an ON pixel which has at least one OFF neighbor; dilation looks for an ON pixel and turns ON all of its immediate neighbors. Hence, small patches of OFF pixels can be closed by dilation followed by erosion (called *closing*) and small patches of ON pixels can be removed by erosion followed by dilation (called *opening*). One can also vary the number of layers (of immediate neighbors) to close bigger holes. The combination used in this work is closing with 3 layers and opening with 2 layers, with TEXTURE being the OFF pixel and SMOOTH being the ON pixel.

2.2. Coefficient Thresholding

The actual denoising is achieved by soft-thresholding the coefficients with thresholds that are spatially and scale-wise adaptive, where the soft-threshold function is defined as $\eta_\lambda(x) = \text{sgn}(x)\max(0, |x| - \lambda)$, and λ is the threshold. For each scale, there are three different thresholds to be used for the different regions, and these values are found empirically. Assume that the noise power σ^2 is known or can be estimated (a reasonable estimate is from the sample variance of the finest scale of the wavelet transform). Since the filters in the wavelet decomposition are known, we can calculate the noise power at scale $s = 2^k$, denoted by σ_k. The thresholds are chosen to be scaled factors of σ_k, and they are $c_E\sigma_k$, $c_T\sigma_k$, and $c_S\sigma_k$ for the EDGE, TEXTURE, and SMOOTH region, respectively. A good set of values are $(c_E, c_T, c_S) = (0.4, 1.4, 1.8)$. Note that especially for large noise σ, it is necessary to soft-threshold the EDGE pixels to avoid ringing.

3. RESULTS

The "cameraman" image is used as the test data, with additive *iid* $N(0, \sigma)$ noise. The wavelets described in [4] are used in 3 scales of nonsubsampled wavelet decomposition. The adaptive threshold scheme is compared with uniform thresholding and Wiener filtering (the version in the MATLAB image processing toolbox). The uniform thresholding method uses the threshold $c_S\sigma_k = 1.8\sigma_k$ for each scale $s = 2^k$, so that it achieves the same smooth background as that obtained by the adaptive method. Table 1 shows the SNR of the various methods for several values of noise strength σ. The SNR for the Wiener filter is taken to be the best one amongst various parameters, but they are not necessarily visually better than the thresholded images.

The denoised images using the adaptive and the uniform thresholding method are shown in Figure 3, for noise $\sigma = 20$. The image using the adaptive scheme (Figure 3(c)) is indeed sharper and retains more details than that using the uniform method (Figure 3(d)), but at the expense of having edges with a slight ringing, a phenomenon more noticeable with increasing σ. To reduce this artifact, one can make the constants c_E closer to c_S and c_T. Figure 2 shows the segmented region. The segmentation method classifies many pixels correctly, but it still needs to be improved and made robust.

σ	Noisy	Adaptive	Uniform	Wiener
10	15.89	20.01	18.03	19.87
15	12.37	17.71	15.99	17.17
20	9.87	16.16	14.64	15.33
25	7.94	14.88	13.52	13.93

Table 1: The SNR (in dB) of the noisy image and of the denoised image using the various methods for several values of σ.

4. CONCLUSION

This work proposed an adaptive thresholding scheme which aims at preserving edges and texture details. This is of course at the expense of significantly more complexity. One can dramatically reduce the complexity by detecting the edges from the original image instead of detecting the multiscale edges in the wavelet domain, and this possibility will be explored in the future, as well as better segmentation methods. Another important issue concerns the values of the thresholds, which need to be derived more systematically. Possible approaches include modeling the different regions with a statistical prior, and finding the Bayesian estimator.

5. REFERENCES

[1] D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol 81, pp. 425-455, 1994.

[2] R.A. Gopinath, M. Lang, H. Guo, and J.E. Odegard, "Enhancement of Decompressed Images at Low Bit Rates," *SPIE* 1994 vol. 2303, pp. 366-377.

[3] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 617-643, March 1992.

[4] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. on PAMI*, vol.14, no.7, pp. 2207-2232, July 1992.

[5] J. Russ, *The Image Processing Handbook*, 2nd ed., IEEE Press, 1994.

[6] D. Wei and C.S. Burrus, "Optimal Wavelet Thresholding for Various Coding Scheme", *Int. Conf. Image Processing* (Washington D.C.), vol. I, pp. 610-613, October 1995.

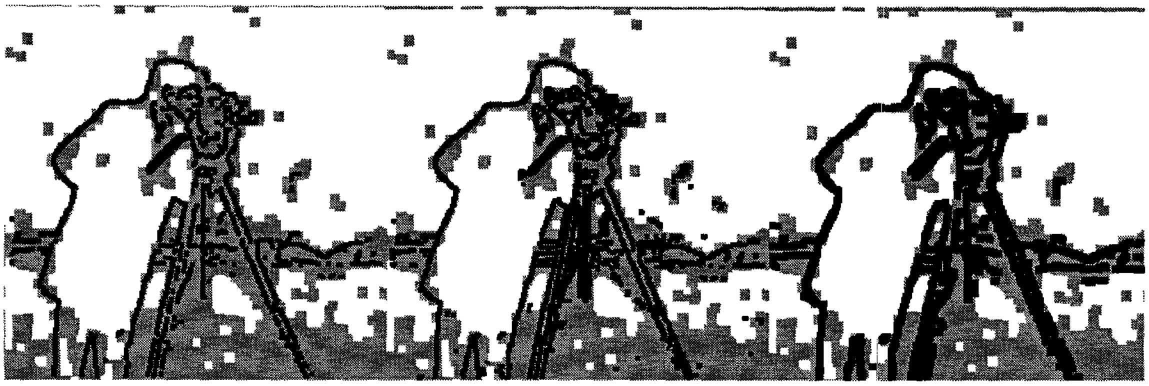

Figure 2: The classification map showing the segmented regions. White is SMOOTH, grey is TEXTURE and black is EDGE.

(a) (b) (c) (d)

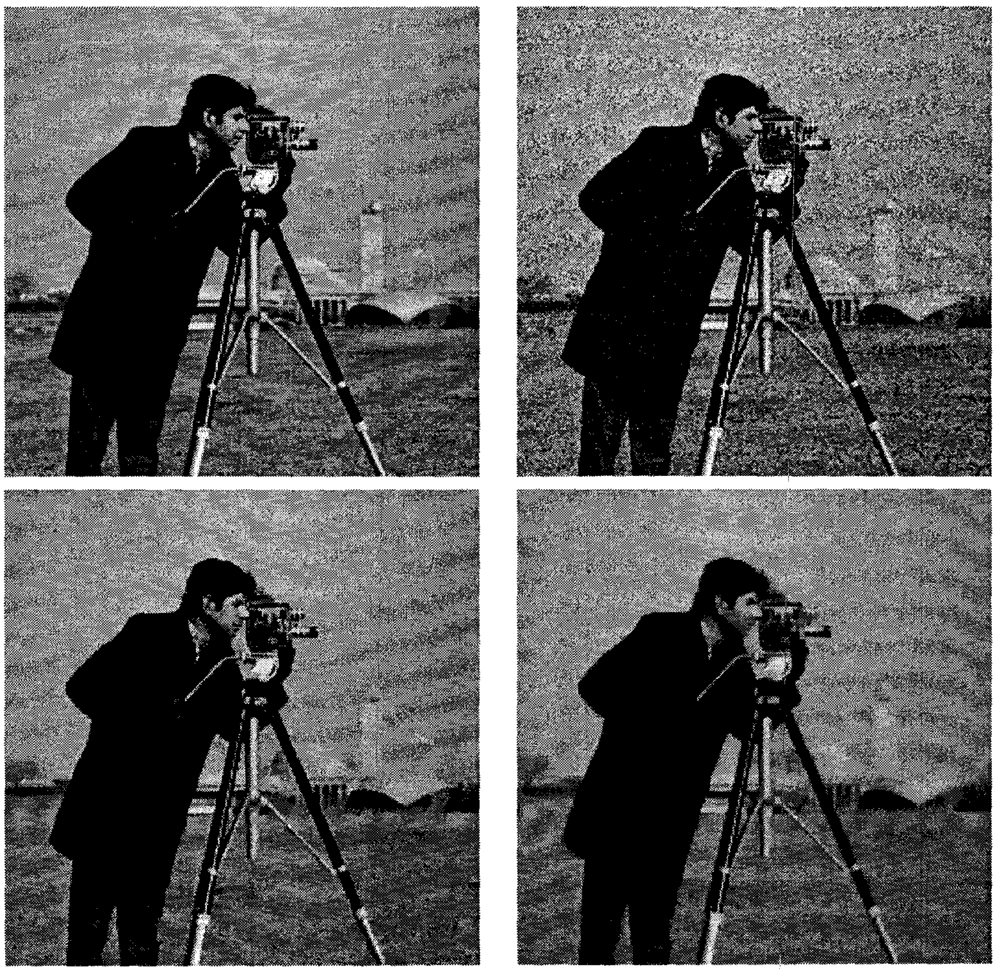

Figure 3: (a) Original; (b) Corrupted with *iid* $N(0, \sigma = 20)$ noise; (c) Denoised using adaptive threshold. (d) Denoised using uniform threshold.